

Потоковый анализ

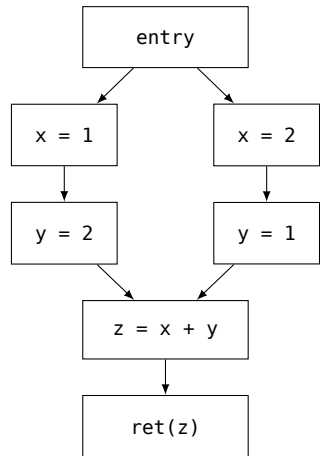
(Data-flow analysis)

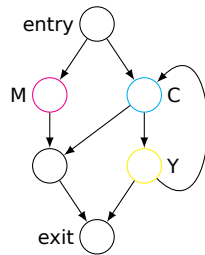
Потоковый анализ

- Статический
- Глобальный (весь CFG)
- Зависит от потока управления
- Вычисление свойств исполнения программы
- Единая формальная модель и теория

Применение

- Reaching definitions (use-def links)
- Live-variable analysis
- Constant propagation
- Constant subexpression elimination
- Dead code elimination





Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



Окружение потокового анализа

- Поточковый граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



Окружение потокового анализа

- Поточковый граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



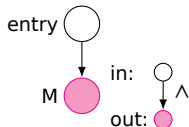
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



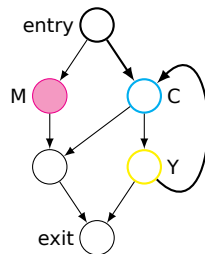
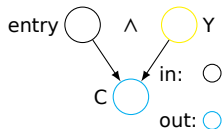
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



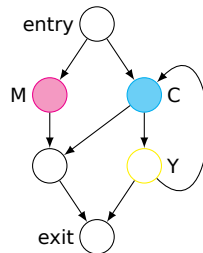
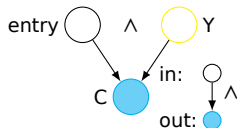
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



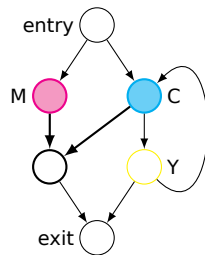
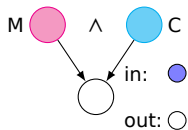
Окружение потокового анализа

- Поточковый граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



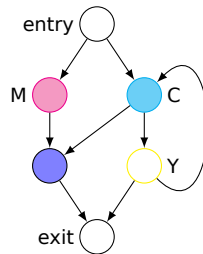
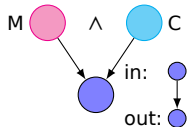
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



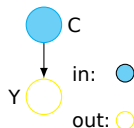
Окружение потокового анализа

- Поточковый граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



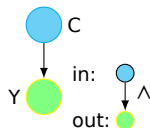
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



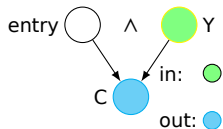
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



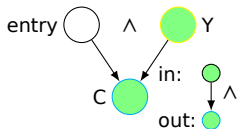
Окружение потокового анализа

- Поточковый граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



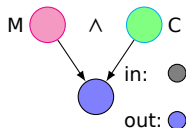
Окружение потокового анализа

- Поточковый граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



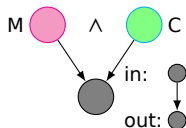
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



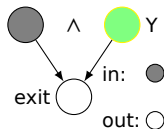
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = T$



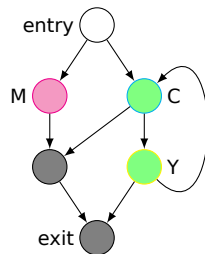
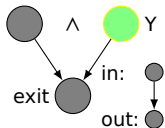
Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



Бинарная операция \wedge (*meet*)

- $x \wedge x = x$ (идемпотентность)
- $x \wedge y = y \wedge x$ (коммутативность)
- $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ (ассоциативность)

Частичный порядок \leq

- $x \leq x$ (рефлексивность)
- $x \leq y \ \& \ y \leq z \Rightarrow x \leq z$ (транзитивность)
- $x \leq y \ \& \ y \leq x \Rightarrow x = y$ (антисимметричность)

Полурешетка $\langle L, \wedge \rangle$ ^{1 2}

- $x \leq y \Leftrightarrow_{def} x \wedge y = x$
- $x < y \Leftrightarrow_{def} x \wedge y = x \ \& \ x \neq y$

¹Выполняются ли свойства частичного порядка при таком определении \leq через \wedge ?

²Можно ли восстановить полурешетку $\langle L, \wedge \rangle$ имея только частичный порядок $\langle L, \leq \rangle$?

Бинарная операция \wedge (*meet*)

- $x \wedge x = x$ (*идемпотентность*)
- $x \wedge y = y \wedge x$ (*коммутативность*)
- $(x \wedge y) \wedge z = x \wedge (y \wedge z)$ (*ассоциативность*)

Частичный порядок \leq

- $x \leq x$ (*рефлексивность*)
- $x \leq y \ \& \ y \leq z \Rightarrow x \leq z$ (*транзитивность*)
- $x \leq y \ \& \ y \leq x \Rightarrow x = y$ (*антисимметричность*)

Полурешетка $\langle L, \wedge \rangle$ ^{1 2}

- $x \leq y \Leftrightarrow_{def} x \wedge y = x$
- $x < y \Leftrightarrow_{def} x \wedge y = x \ \& \ x \neq y$

Свойства полурешеток

Ограниченность снизу

$$\exists \perp \in L : \forall x \in L : \perp \wedge x = \perp \ (\perp \leq x)$$

Ограниченность сверху

$$\exists \top \in L : \forall x \in L : \top \wedge x = x \ (x \leq \top)$$

Высота полурешетки

$$H_L = \max\{|x_1 > x_2 > \dots \in L|\}$$

Обрыв убывающих цепей

$$\forall x_1 > x_2 > \dots \in L : \exists k : \nexists y \in L : x_k > y$$

Произведение полурешеток

$$\begin{aligned} \langle A, \wedge_A \rangle \times \langle B, \wedge_B \rangle &= \langle A \times B, \wedge \rangle, \\ (a, b) \wedge (a', b') &= (a \wedge_A a', b \wedge_B b') \end{aligned}$$

¹Выполняются ли свойства частичного порядка при таком определении \leq через \wedge ?

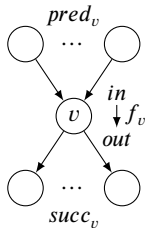
²Можно ли восстановить полурешетку $\langle L, \wedge \rangle$ имея только частичный порядок $\langle L, \leq \rangle$?

- Множество подмножеств S
 $L = 2^S, \wedge = \cap$ (или \cup)
- Натуральные числа
 $L = \mathbb{N}, x \wedge y = \min(x, y)$
- Константные целочисленные значения
 $L = \mathbb{Z} \cup \{T, \perp\}, \perp < \mathbb{Z} < T$
- Иерархия типов в программе
 $L = Types, x \leq y = x <: y$ (*subtype*)

Задача потокового анализа

Окружение потокового анализа

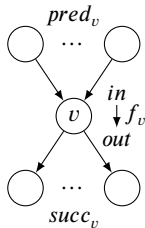
- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



Задача потокового анализа

Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$



Система потоковых уравнений

$D = \downarrow$	$D = \uparrow$
$in_0(v) = out_0(v) = \top$	$in_0(v) = out_0(v) = \top$
$in_i(v) = \bigwedge_{x \in pred_v} out_i(x)$	$out_i(v) = \bigwedge_{x \in succ_v} in_i(x)$
$out_i(v) = f_v(in_i(v))$	$in_i(v) = f_v(out_i(v))$

Maximum Fixed Point (MFP)

Наибольшее решение среди всех решений S

$$out_S(v) \leq out_{MFP}(v) \quad | \quad in_S(v) \leq in_{MFP}(v)$$

Условия сходимости

- Монотонность преобразователей f_v
- Полурешетка $\langle L, \wedge \rangle$ с обрывом цепей

Задача потокового анализа

Окружение потокового анализа

- Поточный граф $G = \langle V, E, v_{entry}, v_{exit} \rangle$
- Направление анализа $D \in \{\downarrow, \uparrow\}$
- Полурешетка свойств $\langle L, \wedge \rangle$ огр. сверху
- Преобразователи свойств $f_{v \in V} : L \rightarrow L$
- Начальная разметка $in_0(v) = out_0(v) = \top$

Преобразователи свойств

Монотонная функция f на $\langle L, \leq \rangle$

$$x \leq y \Rightarrow f(x) \leq f(y)$$

Монотонная функция f на $\langle L, \wedge \rangle$ ³

$$f(x \wedge y) \leq f(x) \wedge f(y)$$

Дистрибутивная функция f на $\langle L, \wedge \rangle$

$$f(x \wedge y) = f(x) \wedge f(y)$$

Система потоковых уравнений

$D = \downarrow$	$D = \uparrow$
$in_0(v) = out_0(v) = \top$	$in_0(v) = out_0(v) = \top$
$in_i(v) = \bigwedge_{x \in pred_v} out_i(x)$	$out_i(v) = \bigwedge_{x \in succ_v} in_i(x)$
$out_i(v) = f_v(in_i(v))$	$in_i(v) = f_v(out_i(v))$

Maximum Fixed Point (MFP)

Наибольшее решение среди всех решений S

$$out_S(v) \leq out_{MFP}(v) \quad | \quad in_S(v) \leq in_{MFP}(v)$$

Условия сходимости

- Монотонность преобразователей f_v
- Полурешетка $\langle L, \wedge \rangle$ с обрывом цепей

³ Докажите эквивалентность определений монотонной функции на $\langle L, \leq \rangle$ и на $\langle L, \wedge \rangle$.

Монотонность преобразователей

$$L = \{T, F\}, F \leq T$$

$$f_{entry} = f_{exit} = id$$

$$f_{loop}(x) = \neg x$$

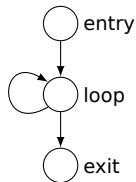
Обрыв убывающих цепей

$$L = \mathbb{R}_0^+ \cup \{T\}, \wedge = min$$

$$f_{entry}(x) = 1$$

$$f_{loop}(x) = x/2$$

$$f_{exit} = id$$



Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP

$$out_{MFP}(v) \leq out_{MOP}(v)$$

⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

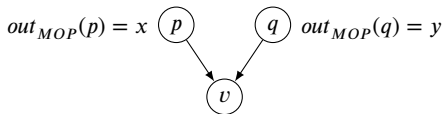
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

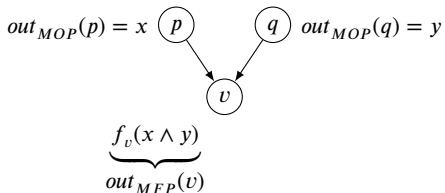
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

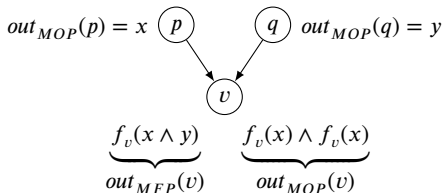
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

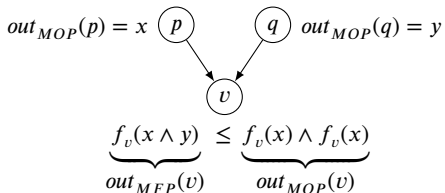
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

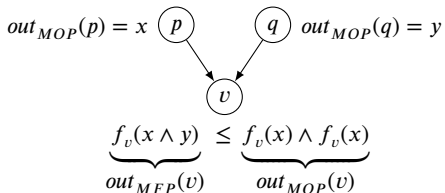
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP ⁵

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

⁵В случае дистрибутивных преобразователей MFP всегда точно — $out_{MFP}(v) = out_{MOP}(v)$.

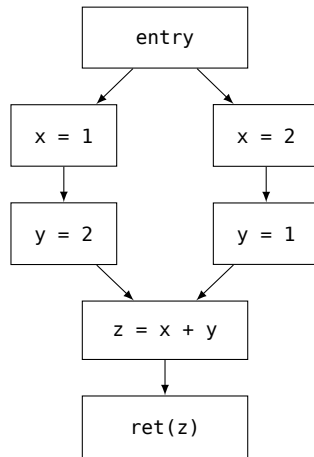
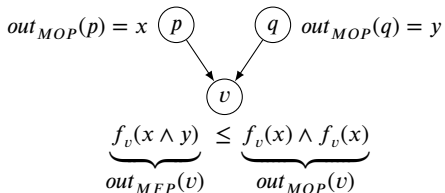
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP ⁵

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

⁵В случае дистрибутивных преобразователей MFP всегда точно — $out_{MFP}(v) = out_{MOP}(v)$.

Точность решения

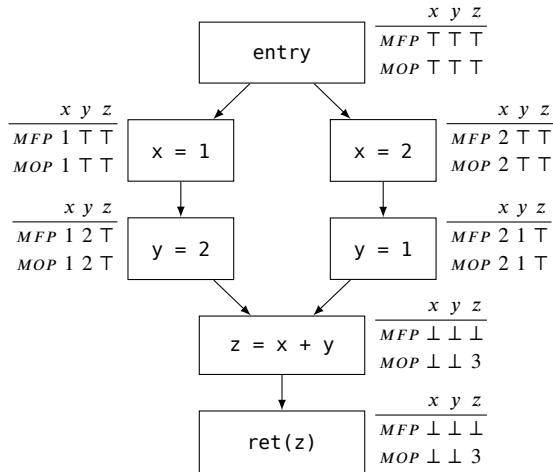
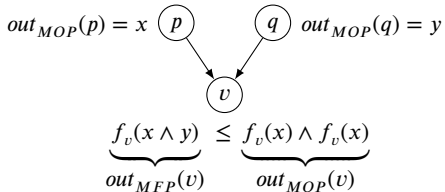
Meet Over Paths (MOP) ⁴

Точное решение по всем путям $v_{entry} \rightarrow \dots \rightarrow v$

$$out_{MOP}(v) = \bigwedge_{v_{entry} \rightarrow \dots \rightarrow v} f_v(\dots (f_{v_{entry}}(\top)) \dots)$$

Безопасность MFP ⁵

$$out_{MFP}(v) \leq out_{MOP}(v)$$



⁴Рассмотрен случай нисходящего анализа $D = \downarrow$, для восходящего $D = \uparrow$ рассуждения аналогичны.

⁵В случае дистрибутивных преобразователей MFP всегда точно — $out_{MFP}(v) = out_{MOP}(v)$.

Control-flow graph

- $CFG = \langle B, E, entry, exit \rangle$
- Каждый блок $b \in B$ содержит одну операцию
- V — множество переменных программы
- $def_v \subseteq B$ — множество *присваиваний* в переменную $v \in V$ (напр. $v = 3$)
- $use_v \subseteq B$ — множество *использований* переменной $v \in V$ (напр. $x = y + v$)

Control-flow graph

- $CFG = \langle B, E, entry, exit \rangle$
- Каждый блок $b \in B$ содержит одну операцию
- V — множество переменных программы
- $def_v \subseteq B$ — множество *присваиваний* в переменную $v \in V$ (напр. $v = 3$)
- $use_v \subseteq B$ — множество *использований* переменной $v \in V$ (напр. $x = y + v$)

Gen-Kill формализм

- $L = 2^S, \wedge = \cup$ или \cap
- $f_b(x) = gen_b \cup (x \setminus kill_b)$
- gen_b — свойства порождаемые блоком b
- $kill_b$ — свойства убиваемые блоком b

Control-flow graph

- $CFG = \langle B, E, entry, exit \rangle$
- Каждый блок $b \in B$ содержит одну операцию
- V — множество переменных программы
- $def_v \subseteq B$ — множество *присваиваний* в переменную $v \in V$ (напр. $v = 3$)
- $use_v \subseteq B$ — множество *использований* переменной $v \in V$ (напр. $x = y + v$)

Gen-Kill формализм

- $L = 2^S, \wedge = \cup$ или \cap
- $f_b(x) = gen_b \cup (x \setminus kill_b)$
- gen_b — свойства порождаемые блоком b
- $kill_b$ — свойства убиваемые блоком b

Следствия

- $\langle L, \wedge \rangle$ — конечная полурешетка
- f_b — дистрибутивные функции⁶
- Анализ *всегда* сходится к точному решению

⁶Докажите дистрибутивность f_b в gen-kill форме.

Control-flow graph

- $CFG = \langle B, E, entry, exit \rangle$
- Каждый блок $b \in B$ содержит одну операцию
- V — множество переменных программы
- $def_v \subseteq B$ — множество *присваиваний* в переменную $v \in V$ (напр. $v = 3$)
- $use_v \subseteq B$ — множество *использований* переменной $v \in V$ (напр. $x = y + v$)

Reaching definitions

$$L = 2^B, \wedge = \cup, D = \downarrow$$

b	$\in def_v$	$\notin def_v$
gen_b	$\{b\}$	\emptyset
$kill_b$	def_v	\emptyset

Live-variable analysis

$$L = 2^V, \wedge = \cup, D = \uparrow$$

gen_b	$\{v \mid b \in use_v\}$
$kill_b$	$\{v \mid b \in def_v\}$

Gen-Kill формализм

- $L = 2^S, \wedge = \cup$ или \cap
- $f_b(x) = gen_b \cup (x \setminus kill_b)$
- gen_b — свойства порождаемые блоком b
- $kill_b$ — свойства убиваемые блоком b

Следствия

- $\langle L, \wedge \rangle$ — конечная полурешетка
- f_b — дистрибутивные функции⁶
- Анализ *всегда* сходится к точному решению

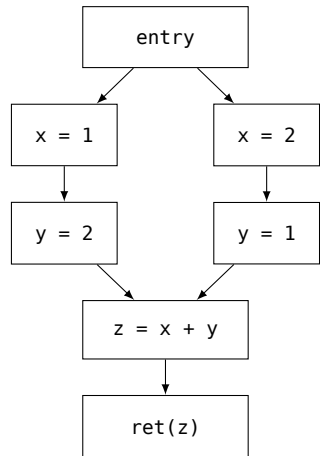
⁶ Докажите дистрибутивность f_b в gen-kill форме.

Достоинства

- Глобальный статический анализ
- Универсальная теоретическая модель
- Простота реализации
- gen-kill формализм гарантирует сходимость и точность

Недостатки

- Результат инвалидируется оптимизациями
- Анализы не комбинируются эффективно
- Не всегда удастся гарантировать сходимость и точность



Спасибо за внимание